Precision, Recall, and F1 Calculation Example

Scenario:
A computer program identifies **7 dogs** in a picture containing **12 actual dogs** and some cats.

Of the 7 dogs identified:
• 4 are actually dogs → **True Positives (TP) = 4**
• 3 are cats → **False Positives (FP) = 3**
The program missed 8 dogs → **False Negatives (FN) = 8**

1■■ Precision (P)
P = TP / (TP + FP) = 4 / (4 + 3) = 4 / 7 = **0.57 (57%)**
Interpretation: When the model predicts "dog," it is correct about 57% of the time.

2■■ Recall (R)
R = TP / (TP + FN) = 4 / (4 + 8) = 4 / 12 = **0.33 (33%)**
Interpretation: The model finds only 33% of all the actual dogs.

3■■ F1 Score
F1 = 2 × (P × R) / (P + R)
F1 = 2 × (0.57 × 0.33) / (0.57 + 0.33) = 2 × 0.1881 / 0.90 = **0.42 (42%)**
Interpretation: The F1 score balances both precision and recall. Here, performance could improve by catching more dogs (increasing recall) without confusing cats for dogs (maintaining precision).

Summary:
Precision = 0.57 (57%)
Recall = 0.33 (33%)
F1 = 0.42 (42%)

Takeaway: This model is fairly cautious but misses many real positives. Improving recall while keeping precision stable would raise the F1 score.